CS229r: IT ∈ CS
LECTURE 16

3/28/2019

TODAY

1. Conclude: Compressing Interactions
2. Direct Product // Amortized Communication.

Review from Last time

1. Theorem: (to be proved): Π can be compressed to length $O(\sqrt{I \cdot k} \log k)$, if $I = IC(\Pi)$, $k = CC(\Pi)$.

2. Ingredient: Correlated Sampling protocol.

   Alice ← P ; Bob ← Q ; P, Q dist on Ω

   Both get R; should output $\omega_A \sim P$, $\omega_B \sim Q$ with $\Pr[\omega_A \neq \omega_B] \leq 2\delta(P,Q)$.

   Protocol: • $R = (a_1, b_1), \ldots (a_i, b_i) \ldots$

   ~~$a_i \in \Omega$, b_i~~ $a_i \sim \mathrm{Unif}(\Omega)$, $b_i \sim \mathrm{Unif}([0,1])$

   • Alice outputs a_i where i is smallest index s.t. $b_i \leq P(a_i)$.

Protocol for Π & Priors

for every node v, at level i in Protocol tree of Π can associate distributions

$$P_v \approx \Pi_i \mid (\Pi_{<i}, X, Y)$$

$$P_v^A \approx \Pi_i \mid (\Pi_{<i}, X)$$

$$P_v^B \approx \Pi_i \mid (\Pi_{<i}, Y)$$

~~Challenge: ...~~

Rough Claim: if $P_v^A \approx \mathrm{Bern}(\frac{1}{2}-\delta)$; $P_v^B = \mathrm{Ben}(\frac{1}{2})$

then $IC = k\delta^2$; $\Pr[\text{Alice + Bob disagree on leaf}]$

$$= O(k\delta)$$

$$= O(\sqrt{k} \cdot \sqrt{I})$$

Next full Protocol + analysis

1. Alice & Bob chose (correlatedly), nodes out of v for every v in tree.
2. find first point v_1 on tree where they disagree.
3. Repeat from v_1 ... until agree on leaf.

Analysis

1. finding first point of disagreement $= O(\log k)$ bits of comm.

   $\Rightarrow$ cost per iteration $= O(\log k)$

2. # iterations = ?

- Let Z_i = indicator variable indicating some iteration starts at level i.

- $\Pr[Z_i = 1] \le \mathop{\mathbb{E}}_{v}\left[\delta(P_v^A, P_v^B)\right]$ where v is random node at level i.

- Cost of i^{th} layer $= V_i$

$$\approx \mathop{\mathbb{E}}_{v}\left[D(P_v^A \| P_v^B)\right]$$

- Rest of analysis: convexity etc.

$$\# \text{ iterations} = \sum_i \mathop{\mathbb{E}}_{v}\left[\delta(P_v^A, P_v^B)\right]$$

$$= \sum_i \mathop{\mathbb{E}}_{v}\left[\sqrt{D(P_v^A \| P_v^B)}\right]$$

$$\le \sqrt{k \cdot \sum_i \mathop{\mathbb{E}}_{v} D(P_v^A \| P_v^B)}$$

$$\le \sqrt{k \cdot I}$$

$\boxtimes$

Direct Product Problems

- Given comm. complexity problem $f : \mathcal{X} \times \mathcal{Y} \to R$
  then n-fold direct product of f is $f^{\otimes n} : \mathcal{X}^n \times \mathcal{Y}^n \to R^n$
  with $f^{\otimes n}((x_1 \dots x_n), (y_1 \dots y_n)) = (f(x_1, y_1), \dots, f(x_n, y_n))$.
- for distributional comm problem, (f, μ), the n-fold product in $(f^{\otimes n}, \mu^{\otimes n})$ $\left[\mu^{\otimes n}(x_1 .. x_n, y_1 .. y_n) = \prod_{i=1}^{n} \mu(x_i, y_i) \right]$.
  - Trivial: $CC(f^{\otimes n}) \leq n \cdot CC(f)$.
  - Question: Can $CC(f^{\otimes n})$ be smaller than $\Omega(n \cdot CC(f))$?
- Note: protocol for $f^{\otimes n}$ doesn't have to work coordinatewise

---

Thm: [BBCR]: $\forall f \quad CC(f^{\otimes n}) \geq \tilde{\Omega}(\sqrt{n} \cdot CC(f))$.

Thm: [BR]: $\forall f \quad CC(f^{\otimes n}) = n \cdot IC(f)$ (*)

Thm: [GKR]: $\exists f$ s.t. $IC(f) = \log(CC(f))$

Corollary: $\exists f$ s.t. $CC(f^{\otimes n}) \leq n \cdot \log CC(f)$

(*) With some careful definitions.

Careful Technical Definition

- How to deal with errors in randomized CC or distribution CC } direct product?
- we expect all coordinates to be right
- if $CC_{\mu,\epsilon}(f) \leq k$ then $CC_{\mu^n, \underbrace{1-(1-\epsilon)^n}_{\text{Not nice!}}}(f^{\otimes n}) \leq k \cdot n$

- Our definition: $\Pi^{(n)}$ solves $f^{\otimes n}$ with error ϵ,
  if $\forall i$ $\boxed{\Pi^{(n)}(\bar{x},\bar{y})_i = f^{\otimes n}(\bar{x},\bar{y})_i \quad \text{w.p.} \geq 1-\epsilon}$

  - Nice: since now ~~the~~ we have natural ϵ-error protocol gives n-fold ϵ-error protocol
  - Not nice: since error of direct. sum $\not\equiv$ error of protocol.

  - But OK since we are mostly proving lower bounds

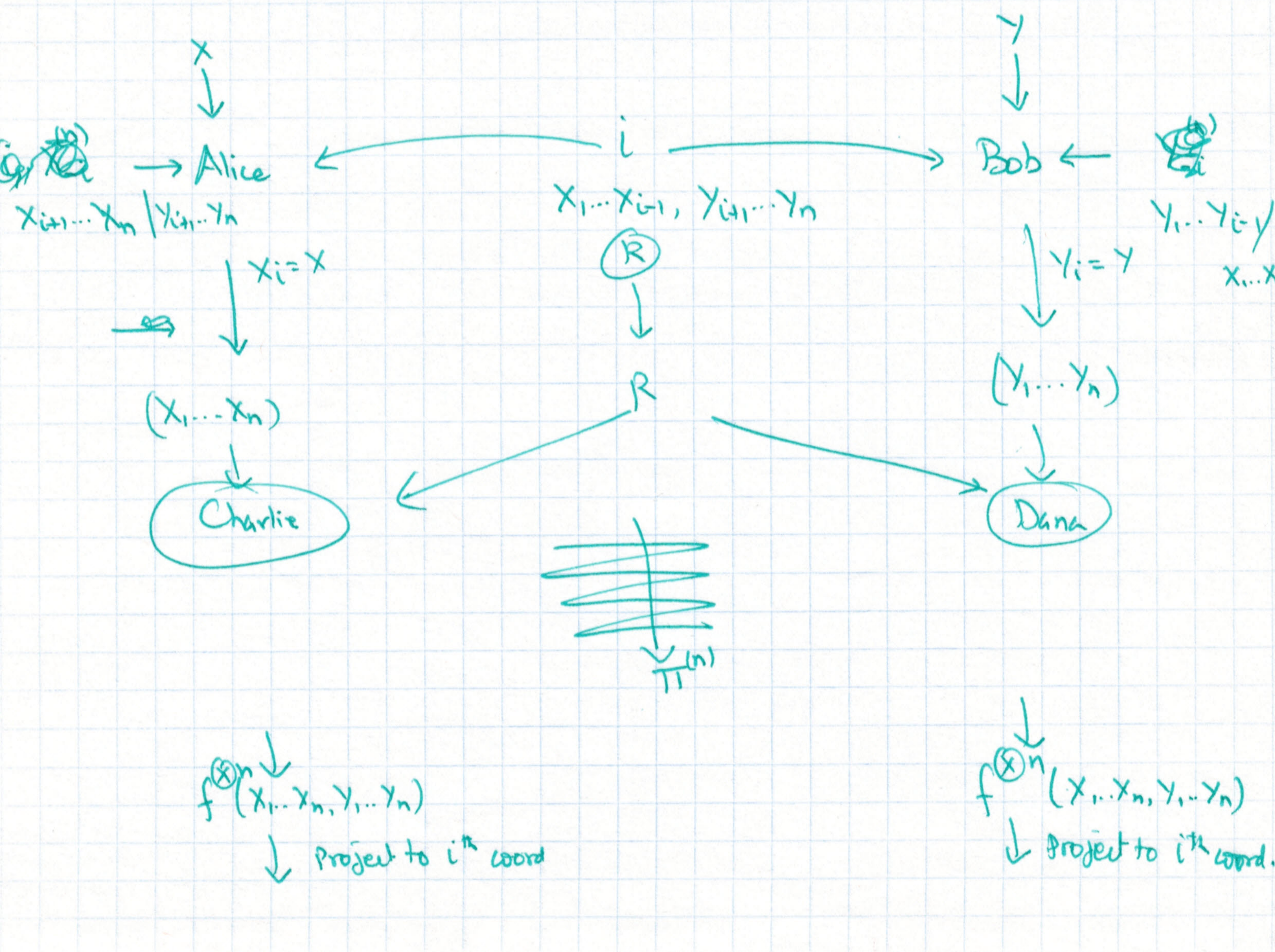
- $CC^n_\epsilon(f) \triangleq \min_{\substack{\bar{\Pi} \text{ that} \\ \text{solves } f^{\otimes n} \\ \text{with } \epsilon\text{-error}}} \{ CC(\bar{\Pi}) \}$; $IC^n_\epsilon(f)$ similar.

Lemma 1: $IC^n_\epsilon(f) = n \cdot IC_\epsilon(f)$

Proof: $\leq$ obvious. Will prove $\geq$. By simulation.

- Suppose $\Pi^{(n)}$ solves $f^{\otimes n}$ with error ϵ & $IC = I$
- Will show how to use $\Pi^{(n)}$ to construct Π that solves f with error ϵ & $IC = \frac{I}{n}$.

Embedding



Claims: 1 Output correct w.p. $1-\epsilon$ ✓ (By Construction)

2 $IC(\pi) = \frac{IC(\pi^{n})}{n}$

$$\frac{1}{n}\sum_{i=1}^{n} I(X_i; \pi^{(n)} \mid X_1 \ldots X_{i-1}, Y_{i+1} \ldots Y_n, i, R; Y_i)$$

$$+ I(Y_i, \pi^{(n)} \mid X_1 \ldots X_{i-1}, Y_{i+1} \ldots Y_n, i, R; X_i)$$

$$\frac{1}{n}\Big(I(X_1 \ldots X_n; \pi^{(n)} \mid R; Y_1 \ldots Y_n) + I(Y_1 \ldots Y_n; \pi^{(n)} \mid R; X_1 \ldots X_n) \Big)$$

Are these equal? Yes! Chain Rule!

$\cancel{X_i - (Y_i) - (X_1 \ldots X_{i-1}, Y_{i+1} \ldots Y_n, i, R, Y_i) - \pi^{(n)}}$

+ the fact that $X_i - (X_1 \ldots X_{i-1}, Y_{i+1} \ldots Y_n, i, R, Y_i) - (\pi^{(n)}; Y_1 \ldots Y_n)$

$$\Rightarrow \sum_{i=1}^{n} I(X_i; \pi^{(n)} \mid X_1 \ldots X_{i-1}, Y_{i+1} \ldots Y_n, i, R; Y_i)$$

$$= \sum_{i=1}^{n} I(X_i; \pi^{(n)} \mid X_1 \ldots X_{i-1}, Y_1 \ldots Y_n, i, R)$$

$$= I(X_1 \ldots X_n; \pi^{(n)} \mid Y_1 \ldots Y_n; R) \qquad \boxtimes$$

Theorem: $CC(f^{\otimes n}) \geq \dfrac{\sqrt{n} \cdot CC(f)}{\log CC(f^{\otimes n})}$.

Proof:
- Take protocol with best comm. complexity C for $f^{\otimes n}$.
- Yield protocols with CC C & IC, $\frac{C}{n}$ for f.
- Compress to get protocol with cc $\log C \cdot \frac{C}{\sqrt{n}}$ for f.

$$\Rightarrow CC(f) \leq \frac{CC(f^{\otimes n})}{\sqrt{n}} \cdot \log C \qquad \boxtimes$$

Concludes BBCR

Towards Braverman-Rao : $CC(f^{\otimes n}) = n \cdot IC(f)$

- since $IC_\epsilon^n(f^{\otimes n}) = n \cdot IC_\epsilon(f)$ & $IC_\epsilon^n(f) \leq CC(f^{\otimes n})$

  we get $CC(f^{\otimes n}) \geq n \cdot IC(f)$.
- So need to show $\leq$. Another "compression" result.
- Central Ingredient: Interactive Correlated Sampling

Alice $\leftarrow P$ $\qquad\qquad$ $Q \longrightarrow$ Bob

$\downarrow$ $\qquad\qquad\qquad\qquad$ $\downarrow$

$X \sim P$ $\qquad\qquad\qquad$ $y = x$ whp $(\geq 1-\epsilon)$

Comm: $D(P\|Q) + 5\sqrt{D(P\|Q)} + O(\log \frac{1}{\epsilon})$

- Contrast with last lecture: 1 No interaction in last lecture.
  2 Here: Interaction + low error (~~and~~ allows ~~of~~ high $D(P\|Q)$).

- ~~Rough Idea~~: Proof: NEXT LECTURE